Graph sparsification

1 Context of the intership

This intership will take place within the ANR project COREGRAPHIE [1] and will be co-supervised by Michel Habib (IRIF - Paris), Mohammed Haddad and Hamida Seba (LIRIS - Lyon). The student will join the GOAL (Graphes Algorithmes et Applications) team at LIRIS (Lyon). This intership can be followed by a PhD thesis.

Contact: {mohammed.haddad, hamida.seba}@univ-lyon1.fr

2 Description

During this intership, we are interested by lossy graph compression and more specifically by graph sparsification which is a powerful method that handles dense graphs. The goal of graph sparsification is to approximate a graph G by a graph G' containing less edges but that retains some properties of the original graph G. The first graph sparsifiers are graph spanners introduced by Chew in [6] and were motivated by the problem of distance similarity of two graphs. A spanner S of a graph G is a sparse graph computed from G in which the shortest-path distance between every pair of vertices is approximately the same in the original graph G as in the spanner S. Graph sparsification is a more general concept that was introduced by [1] to speed-up algorithms computing graph cuts. Given any weighted undirected graph $G = (V, E)$, the authors showed that one could construct a new graph $G_\varepsilon = (V, E_\varepsilon \subseteq E)$, $0 < \varepsilon < 1$, with $|E_\varepsilon| = O(n \log n/\varepsilon^2)$ edges such that the value of every cut in G is within a multiplicative factor of $1 \pm \varepsilon$ of its value in G_ε.

Nowdays, the original concept knows several refinements [8] and extensions to other graph problems such as spectral similarity of graph Laplacians [9], determinant-preservation of matrices [3], to reduce the number of constraints in the binary constraint satisfaction problem (CSP) [2], etc. The concept of vertex sparsification has also been introduced in [5] to approximate the terminal cuts in a graph independently from its size.

In this work, we tackle the problem of graph sparsification under the constraint of preserving low dimensional node representations called *node embeddings*, or at least the induced similarity between them. Mainly we focus on embeddings obtained by walks on graphs such as node2vec [4] and deepwalk [7]. These kind of algorithms use random or biased walks to collect local structures in the graph and nodes' representations are function of the walks rooted at the node using various graph parsing algorithms. Our first work on graph sparsifiers showed promising results. In addition, we also aim at graph sparsifiers preserving graph edit distance which defines the similarity between two graph as the number of edit operations transforming one graph to the other.

[1] `https://coregraphie.projet.liris.cnrs.fr`

References

[1] András A. Benczúr and David R. Karger. Approximating s-t minimum cuts in $o(n^2)$ time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, page 47–55, New York, NY, USA, 1996. Association for Computing Machinery.

[2] Silvia Butti and Stanislav Živný. Sparsification of binary csps. *SIAM Journal on Discrete Mathematics*, 34(1):825–842, 2020.

[3] David Durfee, John Peebles, Richard Peng, and Anup B. Rao. Determinant-preserving sparsification of sddm matrices. *SIAM Journal on Computing*, 49(4):FOCS17–350–FOCS17–408, 2020.

[4] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks, 2016.

[5] Ankur Moitra. Vertex sparsification and oblivious reductions. *SIAM Journal on Computing*, 42(6):2400–2423, 2013.

[6] L. Paul Chew. There are planar graphs almost as good as the complete graph. *Journal of Computer and System Sciences*, 39(2):205 – 219, 1989.

[7] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, page 701–710, New York, NY, USA, 2014. Association for Computing Machinery.

[8] Daniel A. Spielman and Nikhil Srivastava. Graph sparsification by effective resistances. *SIAM Journal on Computing*, 40(6):1913–1926, 2011.

[9] Daniel A. Spielman and Shang-Hua Teng. Spectral sparsification of graphs. *SIAM Journal on Computing*, 40(4):981–1025, 2011.